

Wideband signal transmission system

The invention relates to a transmission system comprising a transmitter for transmitting an input signal to a receiver via a transmission channel, the transmitter comprising a splitter for splitting up the input signal into at least first and second frequency band signals, the transmitter further comprising a first encoder for encoding the first frequency band signal into a first encoded frequency band signal and a second encoder for encoding the second frequency band signal into a second encoded frequency band signal, the transmitter being arranged for transmitting the first and second encoded frequency band signals via the transmission channel to the receiver, the receiver comprising a first decoder for decoding the first encoded frequency band signal into a first decoded frequency band signal and a second decoder for decoding the second encoded frequency band signal into a second decoded frequency band signal, the receiver further comprising a combiner for combining the first and second decoded frequency band signals into an output signal, the receiver further comprising reconstruction means for reconstructing the second decoded frequency band signal when the second decoded frequency band signal is not available.

The invention further relates to a receiver for receiving, via a transmission channel, first and second encoded frequency band signals from a transmitter, to a method of transmitting an input signal via a transmission channel, to a method of receiving, via a transmission channel, first and second encoded frequency band signals and to a speech decoder for decoding first and second encoded frequency band speech signals.

20

A transmission system according to the preamble is known from the paper "An embedded sinusoidal transform codec with measured phases and sampling rate scalability" by Gerard Aguilar et. al. in the proceedings of the 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing, Istanbul, Turkey, 5-9 June 2000, Volume II, pp. 1141-1144.

Such transmission systems may for example be used for transmission of speech signals or audio signals via a transmission medium such as a radio channel, a coaxial cable or an optical fibre. Such transmission systems can also be used for recording of speech

signals on a recording medium such as a magnetic tape or disc. Possible applications are mobile phones, voice over IP (Internet) communication, automatic answering machines and dictating machines.

Figure 1 shows a block diagram of the known transmission system, which

5 transmission system is a so-called scalable wideband speech transmission system. This transmission system comprises a transmitter 12 and a receiver 14. The transmitter 12 and the receiver 14 are coupled via a transmission channel 16. An input speech signal that is supplied to an input 18 of the transmitter 12 is split up into first and second frequency band signals (i.e. spectral portions) by means of a splitter 20. The transmitter 12 further comprises first

10 and second encoders 22 and 24 for encoding the first and second frequency band signals into first and second encoded frequency band signals. These first and second encoded frequency band signals are multiplexed by a multiplexer 23 into a multiplexed signal, which multiplexed signal (carrying the first and second encoded frequency band signals) is transmitted by the transmitter 12 via the transmission channel 16 to the receiver 14. The

15 receiver 14 comprises a speech decoder 60 having a demultiplexer 25 for demultiplexing the multiplexed signal into the first and second encoded frequency band signals and first and second decoders 26 and 28 for decoding the first and second encoded frequency band signals into first and second decoded frequency band signals. The speech decoder 60 further comprises a combiner 30 for combining the first and second decoded frequency band signals into an output signal which is supplied to an output 32 of the receiver 14. Preferably, the first and second encoders 22 and 24 and the first and second decoders 26 and 28 are specifically designed for encoding and decoding the first and second frequency band signals. For

20 example, the first frequency band signal may be a so-called narrowband speech signal having a frequency range of 50-4000 Hz and the second frequency band signal may be a so-called highband speech signal having a frequency range of 4000-7000 Hz. The narrowband speech signal may be encoded and decoded by dedicated narrowband speech coders and decoders. Similarly, the highband speech signal may be encoded and decoded by dedicated highband speech coders and decoders. The decoded narrowband and highband speech signals are combined by the combiner 30 into a so-called wideband speech signal with a frequency range

25 of 50-7000 Hz.

An advantage of such a transmission system is that the narrowband signal can be decoded regardless of the highband signal. Normally both the narrowband and the highband signals are received by the receiver 14 and the speech decoder 60 is able to produce a high quality wideband speech output signal with a frequency range of 50-7000 Hz.

However, when the transmission channel 16 is congested it might occur that frames of the highband signal are not received or are not received correctly by the receiver 14. In such a case the speech decoder 60 is still able to decode the corresponding frames of the narrowband signal and to produce a lower quality narrowband speech output signal with a frequency range of 50-4000 Hz.

The event that a certain frame is not received or is incorrectly received is called a frame erasure. It may be desirable for a transmission system to gracefully handle such frame erasures. In the known transmission system frame erasures are handled either by time scaling (i.e. compressing or expanding in the time domain) the received frames adjacent to the erased frame, or by extrapolating certain parameters of the most recently received frame.

The handling of frame erasures in the known transmission system suffers from a number of drawbacks which negatively influence the quality of the reconstructed speech signal. The handling of frame erasures by time scaling the received frames adjacent to the erased frames is relatively complex and, more importantly, it introduces extra delays because later received frames have to be manipulated in order to correct the erased frames. Furthermore, the handling of frame erasures by extrapolating the parameters of the most recently received frame doesn't always produce the desired result. For instance, for an erased frame which corresponds to the beginning of a new sound it is not possible to reconstruct a similar frame based on the parameters of the last received frame (which corresponds to a different sound).

It is an object of the invention to provide a transmission system as described in the opening paragraph which does not suffer from these drawbacks. This object is achieved in the transmission system according to the invention, which transmission system is characterised in that the reconstruction means are arranged for reconstructing the second decoded frequency band signal from the first decoded frequency band signal. By reconstructing the second decoded frequency band signal from the first decoded frequency band signal, i.e. on the basis of the first decoded frequency band signal, it is possible to avoid the delays which are involved with the time scaling method. Furthermore, the transmission system according to the invention does not suffer from the disadvantage of the method of handling frame erasures by extrapolating certain parameters of the most recently received frame, which method does not produce a correct result when the erased frame corresponds to

the beginning of a new sound. This disadvantage is avoided by reconstructing a frame of the second decoded frequency band signal corresponding to a certain sound from a frame of the first decoded frequency band signal corresponding to the same sound.

An embodiment of the transmission system according to the invention is

- 5 characterised in that the reconstruction means are arranged for reconstructing the second decoded frequency band signal from the first decoded frequency band signal by extending a bandwidth of the first decoded frequency band signal. By means of bandwidth extension a signal with a relatively narrow frequency band can be extended into a signal with a relatively wide frequency band. Several techniques for extending the bandwidth of narrowband signal
10 are known from the paper "A new technique for wideband enhancement of coded narrowband speech", IEEE Speech Coding Workshop 1999, June 20-23, 1999, Porvoo, Finland. These techniques are used to improve the speech quality in a narrowband network.
15 An advantage of the present embodiment of the transmission system according to the invention is that bandwidth extension is a computationally efficient way to reconstruct the second decoded frequency band signal from the first decoded frequency band signal. Moreover, by applying bandwidth extension a very good reconstruction of the erased frames of the second decoded frequency band signal can be obtained, which reconstruction is preferable over simply muting the second decoded frequency band signal.

An embodiment of the transmission system according to the invention is

- 20 characterised in that the reconstruction means are arranged for reconstructing a present frame of the second decoded frequency band signal from a present frame of the first decoded frequency band signal and from a previous frame of the second decoded frequency band signal. By reconstructing a present frame of the second decoded frequency band signal on the basis of a present frame of the first decoded frequency band signal (for instance by means of
25 bandwidth extension) and on the basis of (parameters of) a previous frame of the second decoded frequency band signal an even better reconstruction can be achieved. In some cases it is not always possible to correctly reconstruct a frame of the second decoded frequency band signal on the basis of only the first decoded frequency band signal. For instance, when bandwidth extension is applied to reconstruct a highband speech signal from a narrowband speech signal it is difficult to distinguish between /s/ and /f/ sounds. The reason for this is that these /s/ and /f/ sounds have spectra which are similar in their narrowband part and which are different in their highband part (see Figure 3 in which graph 70 shows the spectrum of the /s/ sound and in which graph 72 shows the spectrum of the /f/ sound). By incorporating
30 information from a previous frame of the second decoded frequency band signal it is clear

which of the /s/ and /f/ sounds was actually included and a correct reconstruction of the second decoded frequency band signal can be made.

A further embodiment of the transmission system according to the invention is characterised in that the first frequency band signal and the first decoded frequency band signal are low frequency band signals and in that the second frequency band signal and the second decoded frequency band signal are high frequency band signals. For example, the transmission system according to the invention can advantageously be used to reconstruct a highband speech signal from a narrowband speech signal.

The above object and features of the present invention will be more apparent 10 from the following description of the preferred embodiments with reference to the drawings, wherein:

Figure 1 shows a block diagram of a prior art transmission system,

15 Figure 2 shows a block diagram of an embodiment of the transmission system according to the invention,

Figure 3 shows spectra of /s/ and /f/ sounds and will be used to explain the operation of the transmission system according to the invention.

In the Figures, identical parts are provided with the same reference numbers.

20 Figure 1 shows a block diagram of a prior art transmission system, which transmission system is a so-called scalable wideband speech transmission system. This transmission system 10 comprises a transmitter 12 and a receiver 14. The transmitter 12 and the receiver 14 are coupled via a transmission channel 16. An input speech signal that is supplied to an input 18 of the transmitter 12 is split up into first and second frequency band signals (i.e. spectral portions) by means of a splitter 20. The transmitter 12 further comprises first and second encoders 22 and 24 for encoding the first and second frequency band signals into first and second encoded frequency band signals. These first and second encoded 25 frequency band signals are multiplexed by a multiplexer 23 into a multiplexed signal, which multiplexed signal (carrying the first and second encoded frequency band signals) is transmitted by the transmitter 12 via the transmission channel 16 to the receiver 14. The receiver 14 comprises a speech decoder 60 having a demultiplexer 25 for demultiplexing the multiplexed signal into the first and second encoded frequency band signals and first and 30 second decoded frequency band signals.

second decoders 26 and 28 for decoding the first and second encoded frequency band signals into first and second decoded frequency band signals. The speech decoder 60 further comprises a combiner 30 for combining the first and second decoded frequency band signals into an output signal which is supplied to an output 32 of the receiver 14. Preferably, the first 5 and second encoders 22 and 24 and the first and second decoders 26 and 28 are specifically designed for encoding and decoding the first and second frequency band signals. For example, the first frequency band signal may be a so-called narrowband speech signal having a frequency range of 50-4000 Hz and the second frequency band signal may be a so-called highband speech signal having a frequency range of 4000-7000 Hz. The narrowband speech 10 signal may be encoded and decoded by dedicated narrowband speech coders and decoders. Similarly, the highband speech signal may be encoded and decoded by dedicated highband speech coders and decoders. The decoded narrowband and highband speech signals are combined by the combiner 30 into a so-called wideband speech signal with a frequency range of 50-7000 Hz.

15 Figure 2 shows a block diagram of an embodiment of the transmission system 10 according to the invention. The transmission system 10 comprises a transmitter 12 for transmitting an input signal to a receiver 14 via a transmission channel 16. The input signal is supplied to an input 18 of the transmitter 12. The transmitter 12 comprises a splitter 20 for splitting up the input signal into a narrowband signal (i.e. the first frequency band signal) and a highband signal (i.e. the second frequency band signal). The splitter 20 comprises a low 20 pass filter 42, a delay element 40 and a subtracter 44. The input signal is supplied to the low pass filter 42 and the delay element 40. The narrowband signal is the result of the filtering of the input signal by the low pass filter 42. The highband signal is the result of subtracting the delayed input signal from the narrowband signal in the subtracter 44. It is important for the 25 low pass filter 42 to have a linear phase characteristic. This may be achieved, for example, by using a finite impulse response filter having a length of 81 so that the filtered signal is delayed by 40 samples. For speech the low pass filter 42 may have a pass band between 0 and 3400 Hz and a stop band between 4000 and 8000 Hz. The delay element 40 is used for compensating the delay that occurs in the low pass filter 42, so that the input signals of the 30 subtracter 44 have a desired phase relation.

 Alternatively, the highband signal may be derived from the input signal by means of a high pass filter (not shown), which is used in stead of the delay element 40 and the subtracter 44.

The narrowband signal is down sampled by a down sampler 46 and applied to a narrowband coder 22 (i.e. the first encoder). This narrowband coder 22 is a coder which is optimized for signals having a narrowband frequency range as described, for example, in ITU standards G.729 or G.728 or in MPEG-4 CELP. The type or operation of this narrowband coder 22 is unimportant to the implementation of the invention. The narrowband coder 22 generates an encoded narrowband signal (i.e. the first encoded frequency band signal). The highband signal is supplied to a highband coder 24 (i.e. the second encoder) for encoding the highband signal into an encoded highband signal (i.e. the second encoded frequency band signal). This highband coder 24 is a coder which is optimized for signals having a highband frequency range as known from, for example, MPEG-4 CELP. The type or operation of this highband coder 24 is unimportant to the implementation of the invention. The encoded narrowband and highband signals are multiplexed in a multiplexer 23 into a multiplexed signal and this multiplexed signal (carrying the encoded narrowband and highband signals) is transmitted by the transmitter 12 via the transmission channel 16 to the receiver 14.

The receiver 14 comprises a speech decoder 60. The speech decoder 60 comprises a demultiplexer 25 for demultiplexing the multiplexed signal into the encoded narrowband and highband signals and a narrowband decoder 26 (i.e. the first decoder) for decoding the encoded narrowband signal and a highband decoder 28 (i.e. the second decoder) for decoding the encoded highband signal. The decoded narrowband signal (i.e. the first decoded frequency band signal) is up sampled by an up sampler 50. In order to filter out undesired highband frequency components which may be introduced in the decoded narrowband signal by the decoder 26 and/or the up sampler 50 the up sampled decoded narrowband signal is filtered by a low pass filter 52. The frequency characteristic of this low pass filter 52 is comparable to the frequency characteristic of the low pass filter 42 in the transmitter 12.

Normally, when a frame of the decoded highband signal is available, the decoded highband signal is supplied via a switch 49 (which is in the lower position) and a delay element 54 to an adder 30 (i.e. the combiner). The decoded narrowband signal (which has been up sampled by the up sampler 50 and filtered by the low pass filter 52) is also supplied to the adder 30. The adder 30 combines the decoded narrowband and highband signals into an output signal which is supplied to an output 32 of the receiver 14. Since it is possible for various signal delays to arise during the decoding of the encoded narrowband and highband signals, the delay element 54 is provided for delaying the decoded highband signal. In the case where the decoded narrowband signal experiences less delay than the

EPO 2002-02-02 10:45:22

decoded highband signal, the delay element 54 may be inserted between the low pass filter 52 and the adder 30.

The speech decoder 60 further comprises a reconstructor 48 (i.e the reconstruction means) for reconstructing (for example by bandwidth extension of the decoded narrowband signal) the decoded highband signal when the decoded highband signal is not available. A frame of the decoded highband signal may not be available, for example, because the corresponding frame of the encoded highband signal was not received at all or because it was not received correctly or because it couldn't be decoded correctly. In such a case the reconstructor 48 reconstructs the missing frame of the decoded highband signal and it is this reconstructed frame that is supplied via the switch 49 (in the upper position) and the delay element 54 to the adder 30. The reconstruction of the missing frame by the reconstructor 48 is done on the basis of a (present) frame of the decoded narrowband signal which is supplied to the reconstructor 48. In addition, the reconstruction of the missing frame may also be (partly) based on (certain parameters of) a previous frame (or previous frames) of the decoded highband signal which is also supplied to the reconstructor 48.

One of the main drawbacks of using bandwidth extension is that there may be multiple extensions of the extended (narrowband) signal. This is very obvious for sounds like /s/ and /f/, whose spectra are similar in their narrowband part and different in the highband part. Figure 3 shows two graphs 70 and 72 illustrating the spectra of these /s/ and /f/ sounds.

In this Figure 3, horizontally the frequency (in Hertz) is plotted, while vertically the amplitude (in dB) of the spectrum is plotted. Graph 70 corresponds to the spectrum of the /s/ sound, while graph 72 corresponds to the spectrum of the /f/ sound. A bandwidth extension system only has the narrowband part available and cannot distinguish between these two sounds. Hence, the extension of the narrowband part of these sounds may lead to audible artefacts.

Suppose that in the transmission system bandwidth extension is used for concealing frame erasures and suppose that the previous frame of data has been received correctly. In the current (present) frame, only the data describing the narrowband part of the wideband signal has been received correctly. The data describing the highband part is lost or erroneous. According to the invention, the highband part can be reconstructed by means of bandwidth extension. This can lead to artefacts as is described above. However, if the highband of the previous frame is received correctly, this can be used to correct some of the errors that are made by extending the bandwidth of the narrowband signal. An important attribute or parameter is the energy of the highband signal. Instead of only using the energy

that is extrapolated from the narrowband signal by the bandwidth extension system, also the energy of the highband from the previously (correctly) received frames can be used. The extrapolated highband signal is then scaled by an averaged value of these energies. For example, if a transmission error in the highband data occurs during an /s/-sound, the

- 5 bandwidth extension system will under-estimate the energy in the highband and as a result it will sound like an /f/. However, if one or more of the previous frames already represented the /s/-sound, the energy of these highband signals can be used to correct the energy level in the highband signal obtained from the bandwidth extension system. This additional information can solve which of the sounds is under consideration. This is an improvement of the system
10 where only bandwidth extension is used for error concealment. In this implementation, the energy of previous frames is used in combination of bandwidth extension. However, also other parameters could be used for this, like for instance the spectral envelope or the pitch period.

Although in the above only a coding scheme having two frequency bands is described, the invention is also applicable to coding schemes having more than two frequency bands. The reconstruction means 48 (with the switch 49) and the speech coder 60 may be implemented by means of digital hardware or by means of software which is executed by a digital signal processor or by a general purpose microprocessor. Furthermore, the reconstruction means 48 may be implemented in the frequency domain or in the time
20 domain.

The scope of the invention is not limited to the embodiments explicitly disclosed. The invention is embodied in each new characteristic and each combination of characteristics. Any reference signs do not limit the scope of the claims. The word "comprising" does not exclude the presence of other elements or steps than those listed in a
25 claim. Use of the word "a" or "an" preceding an element does not exclude the presence of a plurality of such elements.